

The Claims

1. (Previously Presented) Software residing at one or more computer systems collectively operable to execute the software, the software comprising:

a plurality of cluster agents, each cluster agent associated with one of a plurality of nodes, each node comprising a switching fabric integrated to a card and at least two processors integrated to the card, the cluster agent operable to determine a status of the associated node; and

a cluster management engine communicably coupled to the plurality of nodes and operable to:

dynamically allocate a particular subset of the plurality of nodes to a particular job selected from a queue comprising a plurality of jobs, the dynamic allocation based on the determined status of each of one or more of the plurality of nodes; and

execute the job using the particular subset.

2. (Previously Presented) The software of Claim 1, wherein the cluster management engine is further operable to determine a topology of the plurality of nodes based, at least in part, on the determined status of the nodes.

3. (Previously Presented) The software of Claim 2, wherein the topology comprises a three dimensional Torus.

4. (Previously Presented) The software of Claim 1, wherein:
each node comprises at least one host channel adapter; and
the cluster management engine is further operable to dynamically allocate a virtual cluster in the plurality of nodes, the particular subset dynamically allocated to the job comprising at least a subset of the virtual cluster.

5. (Previously Presented) The software of Claim 4, wherein the cluster management engine is further operable to:

dynamically allocate a second particular subset of nodes in the virtual cluster; and
execute a second job using the second particular subset.

6. (Previously Presented) The software of Claim 4, wherein the virtual cluster is associated with a user group.

7. (Previously Presented) The software of Claim 6, wherein the cluster management engine is further operable to verify a user submitting the job based, at least in part, on the user group.

8. (Previously Presented) The software of Claim 4, wherein the cluster management engine is further operable to dynamically allocate a second virtual cluster in the plurality of nodes.

9. (Previously Presented) The software of Claim 1, wherein:
the queue comprises a plurality of jobs awaiting execution, each job submitted by a respective user; and
the queue is one of a plurality of queues, each queue associated with a respective virtual cluster of nodes.

10. (Previously Presented) The software of Claim 1, wherein, to execute the job using the particular subset, the cluster management engine is operable to:
receive a job request comprising one or more job parameters;
determine dimensions of the job based, at least in part, on the one or more job parameters;
dynamically allocate the particular subset based, at least in part, on the determined dimensions; and
execute the job using the particular subset.

11. (Previously Presented) The software of Claim 10, wherein the cluster management engine is further operable to:
select a policy based on the job request; and
dynamically determine the dimensions of the job further based on the selected policy.

12. (Previously Presented) A method comprising:
determining a status of each of at least a subset of a plurality of nodes, each node comprising a switching fabric integrated to a card and at least two processors integrated to the card;

dynamically allocating a particular subset of the plurality of nodes to a particular job selected from a queue comprising a plurality of jobs, the dynamic allocation based on the determined status of each of one or more of the plurality of nodes; and

executing the job using the particular subset.

13. (Previously Presented) The method of Claim 12, further comprising determining a topology of the plurality of nodes based, at least in part, on the determined status of the nodes.

14. (Previously Presented) The method of Claim 13, wherein the topology comprises a three dimensional Torus.

15. (Previously Presented) The method of Claim 12, further comprising dynamically allocating a virtual cluster in the plurality of nodes, the particular subset dynamically allocated to the job comprising at least a subset of the virtual cluster.

16. (Previously Presented) The method of Claim 15, further comprising:
dynamically allocating a second particular subset of nodes in the virtual cluster; and
executing a second job using the second particular subset.

17. (Previously Presented) The method of Claim 15, wherein the virtual cluster is associated with a user group.

18. (Previously Presented) The method of Claim 17, further comprising verifying a user submitting the job based, at least in part, on the user group.

19. (Previously Presented) The method of Claim 15, further comprising dynamically allocating a second virtual cluster in the plurality of nodes.

20. (Previously Presented) The method of Claim 19, wherein the second virtual cluster comprises different nodes from the first virtual cluster.

21. (Previously Presented) The method of Claim 12, wherein executing the job using the particular subset comprises:

- receiving a job request comprising one or more job parameters;
- determining dimensions of the job based, at least in part, on the one or more job parameters;
- dynamically allocating the particular subset based, at least in part, on the determined dimensions; and
- executing the job using the particular subset.

22. (Previously Presented) The method of Claim 21, further comprising:
selecting a policy based on the job request; and
dynamically determining the dimensions of the job further based on the selected policy.

23. (Previously Presented) A system comprising:
a plurality of computing nodes, each computing node comprising a switching fabric integrated to a card and at least two processors integrated to the card, the computing node further comprising an agent operable to determine a status of the computing node and communicate the status to a management node;
the management node, communicably coupled to the plurality of computing nodes and operable to:
dynamically allocate a particular subset of the plurality of computing nodes to a particular job selected from a queue comprising a plurality of jobs, the dynamic allocation based on the determined status of each of one or more of the plurality of computing nodes;
and
execute the job using the particular subset.

24. (Previously Presented) The system of Claim 23, wherein the management node is further operable to determine a topology of the plurality of nodes based, at least in part, on receiving the determined status from the computing nodes.

25. (Previously Presented) The system of Claim 24, wherein the topology comprises a three dimensional Torus.

26. (Previously Presented) The system of Claim 23, wherein the management node is further operable to dynamically allocate a virtual cluster in the plurality of computing nodes, the particular subset dynamically allocated to the job comprising at least a subset of the virtual cluster.

27. (Previously Presented) The system of Claim 26, wherein the management node is further operable to:

dynamically allocate a second subset of computing nodes in the virtual cluster; and
execute a second job using the second subset.

28. (Previously Presented) The system of Claim 26, wherein the virtual cluster is associated with a user group.

29. (Previously Presented) The system of Claim 28, wherein the management node is further operable to verify a user submitting the job based, at least in part, on the user group.

30. (Previously Presented) The system of Claim 26, wherein the management node is further operable to dynamically allocate a second virtual cluster in the plurality of computing nodes.

31. (Previously Presented) The system of Claim 30, wherein the second virtual cluster comprises different computing nodes from the first virtual cluster.

32. (Previously Presented) The system of Claim 23, wherein, to execute the job using the particular subset, the management node is operable to:

receive a job request comprising one or more job parameters;

determine dimensions of the job based, at least in part, on the one or more job parameters;

dynamically allocate the particular subset based, at least in part, on the determined dimensions; and

execute the job using the particular subset.

33. (Previously Presented) The system of Claim 32, wherein the management node is further operable to:

select a policy based on the job request; and

dynamically determine the dimensions of the job further based on the selected policy.

34. (Previously Presented) The software of Claim 1, wherein the card is a motherboard and each node comprises at least one field programmable gate array.

35. (Previously Presented) The method of Claim 12, wherein the card is a motherboard.

36. (Previously Presented) The system of Claim 23, wherein the card is a motherboard.

37. (Previously Presented) Software residing at one or more computer systems collectively operable to execute the software, the software comprising:

a plurality of cluster agents, each cluster agent associated with one of a plurality of nodes, the cluster agent operable to determine a status of the associated node, each node comprising:

at least two first processors integrated to a first card and operable to communicate with each other via a direct link between them; and

a first switch integrated to the first card, the first processors communicably coupled to the first switch, the first switch operable to communicably couple the first processors to six or more second cards each comprising at least two second processors integrated to the second card and a second switch integrated to the second card operable to communicably couple the second processors to the first card and at least five third cards each comprising at least two third processors integrated to the third card and a third switch integrated to the third card;

the first processors being operable to communicate with particular second processors on a particular second card via the first switch and the second switch on the particular second card;

the first processors being operable to communicate with particular third processors on a particular third card via the first switch, a particular second switch on a particular second card between the first card and the particular third card, and the third switch on the particular third card without communicating via either second processor on the particular second card; and

a cluster management engine communicably coupled to the plurality of nodes and operable to:

dynamically allocate a particular subset of the plurality of nodes to a particular job selected from a queue comprising a plurality of jobs, the dynamic allocation based on the determined status of each of one or more of the plurality of nodes; and

execute the job using the particular subset.

38. (Previously Presented) A method comprising:

determining a status of each of at least a subset of a plurality of nodes, each node comprising:

- at least two first processors integrated to a first card and operable to communicate with each other via a direct link between them; and
- a first switch integrated to the first card, the first processors communicably coupled to the first switch, the first switch operable to communicably couple the first processors to six or more second cards each comprising at least two second processors integrated to the second card and a second switch integrated to the second card operable to communicably couple the second processors to the first card and at least five third cards each comprising at least two third processors integrated to the third card and a third switch integrated to the third card;

the first processors being operable to communicate with particular second processors on a particular second card via the first switch and the second switch on the particular second card;

the first processors being operable to communicate with particular third processors on a particular third card via the first switch, a particular second switch on a particular second card between the first card and the particular third card, and the third switch on the particular third card without communicating via either second processor on the particular second card;

dynamically allocating a particular subset of the plurality of nodes to a particular job selected from a queue comprising a plurality of jobs, the dynamic allocation based on the determined status of each of one or more of the plurality of nodes; and

executing the job using the particular subset.

39. (Previously Presented) A system comprising:

a plurality of computing nodes, each computing node comprising:

an agent operable to determine a status of the computing node and communicate the status to a management node;

at least two first processors integrated to a first card and operable to communicate with each other via a direct link between them; and

a first switch integrated to the first card, the first processors communicably coupled to the first switch, the first switch operable to communicably couple the first processors to six or more second cards each comprising at least two second processors integrated to the second card and a second switch integrated to the second card operable to communicably couple the second processors to the first card and at least five third cards each comprising at least two third processors integrated to the third card and a third switch integrated to the third card;

the first processors being operable to communicate with particular second processors on a particular second card via the first switch and the second switch on the particular second card;

the first processors being operable to communicate with particular third processors on a particular third card via the first switch, a particular second switch on a particular second card between the first card and the particular third card, and the third switch on the particular third card without communicating via either second processor on the particular second card; and

the management node, communicably coupled to the plurality of computing nodes and operable to:

dynamically allocate a particular subset of the plurality of computing nodes to a particular job selected from a queue comprising a plurality of jobs, the dynamic allocation based on the determined status of each of one or more of the plurality of computing nodes; and

execute the job using the particular subset.